Introduction to the Adjoint Method

Mohamed Kamal AbdElrahman
s-mohamed.abdelrahman@zewilcity.edu.eg

University of Science and Technology

Zewail City

Outline

Adjoint Sensitivity Analysis of Linear Systems

- Consider the Ordinary Differential Equation (ODE)

$$a(t,z)\frac{d\psi}{dt} + b(t,z)\psi = s(t) \tag{1.1}$$

- z is a design parameter that we want to change in order to minimize the objective function G

$$G = \int_0^T g(\psi, z)\, dt \tag{1.2}$$

Adjoint Sensitivity Analysis of Linear Systems

- To minimize G, we need the gradient with respect to the design parameter z.
- The design parameter z can then be updated with the steepest descent
$$z = z - \alpha \frac{dG}{dz} \tag{1.3}$$
- Local descent involves iteratively determining a descent direction ($\frac{dG}{dz}$) and then taking a step α in that direction and repeating that process until convergence or some termination condition is met

Adjoint Sensitivity Analysis of Linear Systems

- Lets modify the objective function G

$$G = \int_0^T g(\psi, z)\, dt + \int_0^T \lambda(t) R\, dt \tag{1.4}$$

- Where the residual $R = 0 = a(t,z)\frac{d\psi}{dt} + b(t,z)\psi - s(t)$
- The equation is still valid and for any $\lambda(t)$, since we are multiplying $\lambda(t)$ by zero and we just added a zero to the whole equation.
- $\lambda(t)$ is called the adjoint variable.

Adjoint Sensitivity Analysis of Linear Systems

- Using the chain rule to obtain the objective function gradient:

$$\frac{dG}{dz} = \int_0^T \frac{\partial g}{\partial z} + \frac{\partial g}{\partial \psi}\frac{\partial \psi}{\partial z}\,dt + \int_0^T \lambda\left[\frac{\partial a}{\partial z}\dot{\psi} + a\frac{\partial \dot{\psi}}{\partial z} + \frac{\partial b}{\partial z}\psi + b\frac{\partial \psi}{\partial z}\right]dt \tag{1.5}$$

- Integrating by parts the term containing $\frac{\partial \dot{\psi}}{\partial z}$

$$\int_0^T \lambda\, a\frac{\partial \dot{\psi}}{\partial z}\,dt = \lambda a\frac{\partial \psi}{\partial z}\Big|_0^T - \int_0^T \frac{\partial \psi}{\partial z}\frac{d(\lambda a)}{dt}dt \tag{1.6}$$

Adjoint Sensitivity Analysis of Linear Systems

- Collecting all the terms containing the integral of $\frac{\partial \psi}{\partial z}$

$$\frac{dG}{dz} = \int_0^T \frac{\partial g}{\partial z} + \lambda \frac{\partial b}{\partial z}\psi + \lambda \frac{\partial a}{\partial z}\dot{\psi}\,dt + \int_0^T -\frac{\partial \psi}{\partial z}\frac{d(\lambda a)}{dt} + \lambda b \frac{\partial \psi}{\partial z} + \frac{\partial g}{\partial \psi}\frac{\partial \psi}{\partial z}\,dt \tag{1.7}$$

$$\frac{dG}{dz} = \int_0^T \frac{\partial g}{\partial z} + \lambda \frac{\partial b}{\partial z}\psi + \lambda \frac{\partial a}{\partial z}\dot{\psi}\,dt + \int_0^T \frac{\partial \psi}{\partial z}\left[-\frac{d(\lambda a)}{dt} + \lambda b + \frac{\partial g}{\partial \psi}\right] dt + \lambda a \frac{\partial \psi}{\partial z}\Big|_0^T \tag{1.8}$$

Adjoint Sensitivity Analysis of Linear Systems

- If the adjoint variable λ satisfied the following equation

$$-\frac{d(\lambda a)}{dt} + \lambda b + \frac{\partial g}{\partial \psi} = 0 \tag{1.9}$$

- Then the sensitivity of the objective function is

$$\frac{dG}{dz} = \int_0^T \frac{\partial g}{\partial z} + \lambda \frac{\partial b}{\partial z} \psi + \lambda \frac{\partial a}{\partial z} \dot{\psi} \, dt + \lambda a \frac{\partial \psi}{\partial z} |_0^T \tag{1.10}$$

Adjoint Sensitivity Analysis of Linear Systems

- $\frac{d\psi}{dz} = 0$ at t = 0, as the initial condition does not depend on the design parameter.
- Imposing the terminal condition $\lambda(T) = 0$

$$\frac{dG}{dz} = \int_0^T \frac{\partial g}{\partial z} + \lambda \frac{\partial b}{\partial z}\psi + \lambda \frac{\partial a}{\partial z}\dot{\psi}\, dt \tag{1.11}$$

Adjoint Sensitivity Analysis of Linear Systems

- To solve for the sensitivity, we carry two simulations
  1. Solve for ψ
  $$a(t,z)\frac{d\psi}{dt} + b(t,z)\psi = s(t)$$
  2. Calculate the source term for the adjoint equation and solve for λ
  $$\frac{d(\lambda a)}{dt} - \lambda b = \frac{\partial g}{\partial \psi}$$
  3. Evaluate the sensitivity
  $$\frac{dG}{dz} = \int_0^T \frac{\partial g}{\partial z} + \lambda \frac{\partial b}{\partial z}\psi + \lambda \frac{\partial a}{\partial z}\dot{\psi}\, dt$$